Examining the designs of computer-based assessment and its impact on student engagement, satisfaction, and pass rates

Journal Item

Computers in Human Behavior 76 (2017) 703–714

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Examining the designs of computer-based assessment and its impact on student engagement, satisfaction, and pass rates

Quan Nguyen [a], Bart Rienties [a, *], Lisette Toetenel [b], Rebecca Ferguson [a], Denise Whitelock [a]

[a] Institute of Educational Technology, The Open University, UK
[b] Bank Julius Baer & Co, Zurich, Switzerland

ARTICLE INFO

Article history:
Received 30 June 2016
Received in revised form
22 February 2017
Accepted 12 March 2017
Available online 17 March 2017

Keywords:
Computer-based assessment
Learning design
Learning analytics
Academic retention
Learner satisfaction
Virtual learning environment

ABSTRACT

Many researchers who study the impact of computer-based assessment (CBA) focus on the affordances or complexities of CBA approaches in comparison to traditional assessment methods. This study examines how CBA approaches were configured within and between modules, and the impact of assessment design on students' engagement, satisfaction, and pass rates. The analysis was conducted using a combination of longitudinal visualisations, correlational analysis, and fixed-effect models on 74 undergraduate modules and their 72,377 students. Our findings indicate that educators designed very different assessment strategies, which significantly influenced student engagement as measured by time spent in the virtual learning environment (VLE). Weekly analyses indicated that assessment activities were balanced with other learning activities, which suggests that educators tended to aim for a consistent workload when designing assessment strategies. Since most of the assessments were computer-based, students spent more time on the VLE during assessment weeks. By controlling for heterogeneity within and between modules, learning design could explain up to 69% of the variability in students' time spent on the VLE. Furthermore, assessment activities were significantly related to pass rates, but no clear relation with satisfaction was found. Our findings highlight the importance of CBA and learning design to how students learn online.

1. Introduction

The field of Learning Design has emerged during this century. It seeks to develop a conceptual and descriptive framework of teaching practices, making use of technology, in order to make learning designs explicit, sharable, and reusable (Conole, 2012; Kirschner, Strijbos, Kreijns, & Beers, 2004). Since assessment is a common element of learning design, it is of great importance to understand how assessment activities are configured, and to represent them in an explicit, sharable, and reusable manner. In terms of capturing the design of assessment activities, an increasing number of educators use computer-based assessment (CBA). CBA can be defined as assessment presented using digital means and submitted electronically. CBA offers many advantages over traditional forms of assessment as it is electronically built, and therefore generates user data on assessment activities.

CBA allows teachers to compare the designs of their assessment approaches across different tasks and modules. For example, Toetenel and Rienties (2016) compared 157 modules at The Open University (OU) and found that, on average, 21.50% of students' total workload was allocated for assessment, although substantial variation (SD = 14.58%, range 0%–78%) was found amongst these modules. By representing CBA in an explicit way, educators can learn from each other by comparing, reusing, and adapting peers' CBA designs in their own learning context.

The impact of CBA designs on the learning processes of students may be better understood with the support of learning analytics, which make use of the digital traces of learners' interactions in a virtual learning environment (VLE) that are preserved in log-files. Recent learning analytics research has found that the way in which teachers design tasks and assessments at a micro level (within one assessment or task: see for example Greiff, Wüstenberg, and Avvisati (2015)) and a macro level (across various assessments within or across modules: see for example

* Corresponding author. Jennie Lee Building, Walton Hall Campus, Milton Keynes, MK7 6AA, UK.
E-mail address: bart.rienties@open.ac.uk (B. Rienties).

http://dx.doi.org/10.1016/j.chb.2017.03.028
0747-5632/

Lockyer, Heathcote, and Dawson (2013)) can influence how students are engaging with CBA tasks and their academic performance (Koedinger, Booth, & Klahr, 2013; Rienties & Toetenel, 2016; Toetenel & Rienties, 2016).

For example, in a fine-grained log file study of one Programme for International Assessment (PISA) CBA task on climate control that compared complex process solving amongst 16,219 children in 44 countries, Greiff et al. (2015) found substantial differences in individual strategies. These strategies had a significant impact on individual performance and also highlighted differences in problem-solving strategies between countries. In a follow-up study of 1476 Finnish children completing nine complex problems, Greiff, Niepel, Scherer, and Martin (2016) found that there was an optimal level of effort spent on these tasks and consecutive performance, as well as a negative relation between the frequency of changes made in learning strategy and performance. These micro-level CBA studies provide rich and complex understandings of how students engage with specific assessment tasks (Koedinger et al., 2013; Vandewaetere, Desmet, & Clarebout, 2011), and may help teachers to make decisions about intervention both on individual tasks and on a broader, module-based level.

Similarly, but at a more macro-level of analysis, Rienties and Toetenel (2016) studied the activity of 111,256 students on 151 modules at the OU using multiple regression models. The study found that learning designs strongly predicted VLE behaviour and performance of students. In follow-up research, Rogaten, Rienties, and Whitelock (2016) compared the learning gains of 17,700 students on 110 Science and Social Science modules and found individual differences in learning gains as well as significant differences in assessment practices within and across OU modules. Module characteristics accounted for 6%–33% of variance in students' initial achievements and 19%–26% of variance in subsequent learning gains. In other words, recent preliminary findings suggest that how teachers design CBA influences how students learn over time on both a micro and a macro level.

By aligning the designs of CBA with fine-grained data relating to students' engagements with the VLE, together with their satisfaction and performance, educators may be equipped with valuable insights as to how their students are "reacting" to the design of CBA (Koedinger et al., 2013; Rienties & Toetenel, 2016; Toetenel & Rienties, 2016). While, to the best of our knowledge, the study by Rienties and Toetenel (2016) was the first to link aggregate learning design data of CBA and other activities on a large scale with actual student behaviour and cognition across a large number of modules, the study did not deal with how educators use CBA on a week-by-week basis. Aligning how educators balance CBA on a weekly basis with what students actually do in the VLE can not only advance our insights into CBA, but may also help to strengthen the links between learning analytics and CBA.

Therefore, firstly we aim to understand how educators design and implement CBA on a weekly basis. Specifically, we will investigate how educators have allocated time for seven types of learning activity (assessment, assimilative, finding information, communication, productive, interactive, and experiential) at both module level (for 74 modules) and weekly level (for 37 modules).

Secondly, by building on previous research (Rienties & Toetenel, 2016; Toetenel & Rienties, 2016), we will investigate the impact of CBA designs on students' behaviours in the VLE using log-files data, satisfaction levels, and module pass rates. This element of the study is of particular interest as a large body of assessment literature has indicated that assessment drives learning (Bearman et al., 2016; Segers, Dochy, & Cascallar, 2003), but limited empirical evidence on a macro level is available to confirm this claim. Our analysis will use a combination of visualisation techniques and fixed effect models on 72,377 registered students studying 74 modules across different disciplines.

Finally, while learning analytics may highlight key correlations and even causation in relation to learning design, CBA, student engagement and learning, it is important to be aware that learning always takes place in a context. Therefore, our third and final aim is to unpack how six teachers designed their modules and understand how their weekly design decisions influenced student learning, using a case-study approach.

2. Computer-based assessment, learning design and learning analytics

CBA has a lot of potential to demonstrate how students learn and solve complex tasks (Greiff et al., 2015; Ras, Whitelock, & Kalz, 2015; Tempelaar, Rienties, & Giesbers, 2015). It offers many advantages when compared to more traditional forms of assessment. In distance-learning settings the most relevant benefits are: speed, cost-reduction (Terzis & Economides, 2011), automatic feedback provision (Ras et al., 2015), record keeping, and more authentic interactive assessment options. These options include intelligent tutoring (Koedinger et al., 2013), adaptive CBA (Tempelaar et al., 2015; Vandewaetere et al., 2011), and authentic virtual labs (Scherer, Meßinger-Koppelt, & Tiemann, 2014).

In a study of six online and two blended courses, Agudo-Peregrina, Iglesias-Pradas, Conde-González, and Hernández-García (2014) found that interactions with CBA tools, interactions with peers and teachers, as well as active participation were significant predictors of academic performance on online courses. Similarly, in a blended course in mathematics and statistics that relied intensively on CBA, Tempelaar et al. (2015) found that behaviour and mastery in CBA assessments were good predictors of academic performance. In contrast, most metrics not linked to CBA that can easily be gathered from a VLE (such as time spent and number of downloads) provided limited insight into student progression.

In a recent study of computer programming courses using CBA design, Brito and de Sá-Soares (2014) found that a high frequency of CBA, at weekly level, was one of the most effective ways of setting students on the route to success in an introductory computer programming course. Weekly CBA may help to speed up the cycle of productive failure (Kapur, 2008) – "fail fast to learn sooner" (Brito & de Sá-Soares, 2014) – because it can provide automated feedback. In a follow-up fine-grained study of 1080 students using worked examples (i.e., accessing the help function of a particular task) during CBA in mathematics, Tempelaar et al., 2017 found that students with sub-optimal learning strategies tended to use these worked examples at the end of their learning process, while students with effective metacognitive strategies used these at the beginning of their learning process. In other words, there is some support for the idea that CBA in conjunction with learning analytics data enables researchers and educators to unpack and predict successful learning.

Most studies that have conceptualised and tested different variations of CBA influence on student engagement and learning outcomes have used a single module (Tempelaar, Rienties, & Nguyen, 2017 ; Tempelaar et al., 2015) or a single unit of assessment context (Greiff et al., 2015, 2016). With the advent of learning analytics (Ferguson & Buckingham Shum, 2012; Rienties & Toetenel, 2016) and learning design (Conole, 2012; Rienties & Toetenel, 2016), researchers can now compare how teachers across a range of modules and disciplines are designing their modules. This could allow us to test which learning designs lead to better student engagement and learning.

2.1. CBA as a part of learning design

Conole (2012, p. 121) described learning design as "a methodology for enabling teachers/designers to make more informed decisions in how they go about designing learning activities and interventions, which is pedagogically informed and makes effective use of appropriate resources and technologies". Learning design focuses on what students do as part of their learning, rather than on the content that is delivered by the teaching. Within this journal, it is increasingly evident that learning design is an important driver for learning (Hernández-Leo, Moreno, Chacón, & Blat, 2014; Moreno-Ger, Burgos, Martínez-Ortiz, Sierra, & Fernández-Manjón, 2008; Rienties & Toetenel, 2016).

For example, in a large scale study of 111,256 students on 151 modules at the OU using multiple regression models, Rienties and Toetenel (2016) found that learning designs strongly predicted VLE behaviour and performance of students. Additionally, Li, Marsh, Rienties, and Whitelock (2016) carried out a study of 422 undergraduate blended and online courses that involved logistical regression modelling of learner satisfaction using 232 variables including learning design, learner characteristics and assessment types. They found that learning design decisions significantly influenced learner satisfaction. While the most important key driver for learner satisfaction was the quality of teaching materials, the second most important driver was the quality of assessment (Q36: Overall, I was satisfied with the assessment on this module). In addition, assignment instructions (Q9: The instructions on how to complete the assignments were clear) and assignment completion (Q10: Taking part in optional exercises or activities to test my understanding helped me to learn) were consistently among the top 10 drivers of learner satisfaction (Li et al., 2016).

These studies linked large datasets across a range of modules in both online and blended learning settings (Arbaugh, 2014; Li et al., 2016; Rienties & Toetenel, 2016). Together, they demonstrated that it is possible to obtain important and generalisable insights that extend beyond the research findings associated with a single module or discipline. This can be done by carrying out a cross-sectional study that analyses the impact of learning design on assessment strategies, learner satisfaction and academic performance across a range of modules.

2.2. Research questions

While substantial advances in our understanding of assessment strategies have recently been made (Bearman et al., 2016; Ras et al., 2015; Tempelaar et al., 2015), few studies have examined effective assessment strategies across a number of modules. Furthermore, in a recent study Rienties and Toetenel (2016) found that although learning design is widely studied as a way to improve course design (Armellini & Aiyegbayo, 2010; Koedinger et al., 2013; MacLean & Scott, 2011), few institutions have captured and updated these data in order to reflect on how these courses are delivered to students.

Research question 1 focuses on how teachers design and use CBA in online distance learning, and how these designs relate to the broader notions of learning design developed by Conole (2012). By addressing research question 2 we aim to understand how CBA decisions made by teachers influence student engagement in the VLE, and how this engagement may influence student satisfaction and pass rates. In comparison to Rienties and Toetenel (2016), a unique new contribution of this study is that we will investigate the impact of weekly CBA design decisions in 37 modules on student engagement in research question 2. This may be conceptually and empirically important, as CBA tasks in, say, week 4 may have an impact not only on student engagement in week 4, but also on engagement before and after the assessment activity. Therefore, using regression and fixed effects models in the first part of this study we will focus on two research questions and build on the work of Rienties and Toetenel (2016) by applying a learning analytics perspective to a large dataset.

1) To what extent do educators use individual assessment designs and how do these relate to learning design principles?
2) To what extent are educators' assessment design decisions associated with student engagement in the VLE, student satisfaction and pass rates?

Two recent meta-analyses on learning analytics (Papamitsiou & Economides, 2014, 2016) and a large-scale review of learning analytics impact studies for the European Commission (Ferguson et al., 2016) indicated that relying extensively on large data and on quantitative proxies of learning can make it difficult to understand the context in which learning and CBA strategies take place. Therefore, research question 3 prompts a more fine-grained analysis of the assessment and learning design strategies of six modules designed for first-year undergraduates.

3) Which, if any, weekly assessment strategies are associated with student engagement over time?

3. Method

3.1. Setting and participants

This study took place at the OU, a distance-learning institution with an open-entry policy and the largest university in the UK. As previous research has found substantial differences between postgraduate and undergraduate learning designs (Li et al., 2016; Rienties & Toetenel, 2016), this study included only undergraduate modules that have run since 2014 with at least 200 registered students per module. In total, 72,377 students were enrolled in these 74 undergraduate module presentations, with an average of 978 (range: 207–3707) registered students. These modules related to various disciplines (25% in Science & Technology, 22% in Arts & Social Sciences, 14% in Business & Law, 9% in Education & Languages, and 30% in other disciplines).

There were more female students (54%) than male students (46%) studying these 74 modules. The majority of these students were from the UK (96%) and declared their ethnicity to be 'white' (87%). Students varied considerably in age, with 25% under 25 years old, 36% aged 26–35, 21% aged 36–45, 12% aged 45–55, and 6% aged 56 and over. More than half of them were working full-time (52%), while 19% were working part-time, 8% were looking after the home/family, and 6% were unemployed and looking for a job. Regarding learners' qualifications, there are no formal academic entry requirements at undergraduate level at the OU. In this study, 41% of the students had A levels or equivalent (suggesting they had two or more years of post-compulsory schooling), 31% had less than A levels (suggesting they had not progressed beyond compulsory schooling), 22% had higher education degrees, and 4% had a postgraduate qualification. On average, 9% of the students had a reported disability.

Longitudinal data relating to learning design were available on a weekly basis for 37 of the 74 module presentations (these modules had 45,190 registered students). While learning designs had originally been modelled for each two- to eight-week block of study, the research findings of (Rienties & Toetenel, 2016; Toetenel & Rienties, 2016) prompted the learning design team to model learning designs on a week-by-week basis.

3.2. Instruments

3.2.1. Learning design mapping

A university-wide learning initiative to use learning design data was developed by Conole (2012) for quality enhancement purposes. This initiative introduced a process of “module mapping” or “coding learning activities”. The taxonomy used to do this had been enhanced by the Jisc-sponsored Open University Learning Design Initiative (OULDI) (Cross, Galley, Brasher, & Weller, 2012), which took place over five years in consultation with eight higher education institutions. This initiative involved analysing and providing visualisations of the learning activities and resources associated with each of the university's modules (Rienties & Toetenel, 2016). For a detailed description of the mapping process and reliability of this approach, we refer to our previous article in this journal (Rienties & Toetenel, 2016).

In contrast to instructional design, learning design is process based (Conole, 2012); following a collaborative design approach in which practitioners make informed design decisions with a pedagogical focus. To do this, they use representations in order to build a shared vision, which is intended to ensure that these practitioners jointly provide a cohesive learning experience for students.

As illustrated in Table 1, *Assessment activities* include all work that is focused on summative, formative or self-assessment. According to the assessment handbook of The Open University UK (2016), common types of assessment include: assignments, oral or practical assessments, projects, examinations, dissertations, and portfolios. Each module usually includes two assessment components. The first is known as continuous assessment because it takes place throughout the module. Continuous assessment includes tutor-marked assignments (TMAs), and computer-marked assignments (CMAs). CMAs may be submitted via online forms or may be interactive (iCMAs). End-of-module assessments (EMAs) are the second assessment component. These involve examinations (usually hand-written), or other pieces of work such as dissertations, projects, or portfolios. As this takes place in a distance-education context, the assessment is mediated by technology and so is considered to be CBA. This definition could be problematic because in some cases learning and assessment materials are available both online and offline, or learners can print the assessment materials and submit by post. However, they meet our definition of CBA because all the assessment materials are provided through digital means and the majority of learners submit their work online.

The *Assimilative activities* in Table 1 relate to tasks in which learners attend to discipline-specific information. These tasks include reading text (online or offline), watching videos and listening to audio files. The next five types of activity can be used to create an interactive environment in which learning with others can take place (Ferguson & Buckingham Shum, 2012). By *Finding and handling information* provided by others, rather than using only content provided by educators, learners are able to take responsibility for their own learning. During *Communicative activities*, students communicate with others about module content (Arbaugh, 2014; Kirschner et al., 2004), and during *Productive activities* they build and co-construct new artefacts. *Experiential activities* provide learners with opportunities to apply their learning in a real-life setting. Finally, *interactive activities* also provide opportunities to apply learning, but in safe settings such as simulations (Scherer et al., 2014).

The seven types of learning activity were measured in terms of the duration in hours that was allocated for each type of activity. This was determined through a detailed and comprehensive the mapping process between learning designers and instructors (see Rienties & Toetenel, 2016 for further details). These measurements were captured at a module level (for 74 modules) and at a weekly level (for 37 modules). The number of credits to be gained determined the total workload of each module, which is the sum of the time allocated for all seven types of learning activity. Generally speaking, each credit is associated with 10 h of study (so 30 credits = 300 h and 60 credits = 600 h). However, the actual workload can be different, and depends on each module's learning design. Of the 74 modules covered by this study, two offered 15 credits, 40 offered 30 credits, and 32 modules offered 60 credits.

3.2.2. VLE data: student engagement

In line with Tempelaar et al. (2015), two different types of data were gathered for each module from the university's Moodle VLE: average time spent on the VLE per week (in minutes); and average time spent per session on the VLE (in minutes). Subsequent derivatives of these two types of data per week were recorded from week -3 until week 40 (data streams typically start three weeks before the official start of the module). More fine-grained tracking data were available for some modules, providing information about types of content, materials and ICT tools (such as wikis, videoconference and discussion forums). However, given the diversity in usage and because not all modules used all the ICT tools, we focused on aggregate user statistics per week across the VLE as our proxy for student engagement.

Table 1
Learning design activities.

	Type of activity	Example	Mapped/Coding
Assessment	All forms of assessment (summative, formative and self assessment)	Write, Present, Report, Demonstrate, Critique.	Based upon estimated average time spent per week, unless specified by educator.
Assimilative	Attending to information	Read, Watch, Listen, Think about, Access.	Based upon word count or time per audio/video clip.
Finding and handling information	Searching for and processing information	List, Analyse, Collate, Plot, Find, Discover, Access, Use, Gather.	Indicated by educator or estimated by Learning Design team.
Communication	Discussing module-related content with at least one other person (student or tutor)	Communicate, Debate, Discuss, Argue, Share, Report, Collaborate, Present, Describe.	
Productive	Actively constructing an artefact	Create, Build, Make, Design, Construct, Contribute, Complete.	
Experiential	Applying learning in a real-world setting	Practice, Apply, Mimic, Experience, Explore, Investigate.	
Interactive/adaptive	Applying learning in a simulated setting	Explore, Experiment, Trial, Improve, Model, Simulate.	

Adapted from (Rienties & Toetenel, 2016)

3.2.3. Learner satisfaction

Since its foundation nearly 50 years ago, the OU has consistently collected learner feedback in order to improve its learning experience and learning designs. The Student Experience on a Module (SEAM) survey is employed as part of this process, just as with other learner satisfaction instruments (Onwuegbuzie et al., 2007; Zerihun, Beishuizen, & Os, 2012). This standard questionnaire is sent to all students who are still registered at the end of the module.

Following our analysis of key drivers of the learning experience of 115,000 students (Li et al., 2016), for this analysis we used the aggregate scores of five core items (out of 40) from the SEAM survey that have been shown to drive learner satisfaction. These five items measured students' satisfaction with regard to (1) teaching materials, (2) assessment on module studied, (3) advice and guidance provided for module study, (4) integration of materials, and (5) career relevance, scaling from one to five in which one means “definitely agree”, and five means “definitely disagree”. On average, 81% of the respondents were satisfied with their learning experience (SD = 11%), with a minimum of 39% and a maximum of 97% (Table 2). The average response rate for 74 modules was 31.67% (SD = 7%, min = 18.71%, max = 53.30%).

3.2.4. Pass rate

On average, there were 978 registered students per each module (SD = 824), with a minimum of 205 and a maximum of 3707. The pass rate was calculated as the percentage of registered students who completed and passed the module (Range: 34%–90%; Mean = 63.57, SD = 10.21).

3.3. Data analysis

All data were collected at an aggregate, module level and were anonymised in line with the university's ethical guidelines. Based on module ID and the year and month in which the module was presented (some modules run twice a year), the learning design data for 74 module presentations were merged with the available VLE and pass rate data, which allowed us to analyse the combined data for 71 modules.

In preparation for the panel analysis of 37 modules (those with weekly longitudinal data), a Hausman test was used to differentiate between a fixed effects and a random effects model. This test checks whether the coefficients estimated by the random effects estimator are the same as the ones estimated by the consistent fixed effects estimator (Hausman, 1978). Our result supported the assumption of correlation between observation errors and predictors. For this reason, a fixed effects model was used as it removes the effect of time-invariant characteristics to assess the net effect of the predictors on the outcome.

Variance inflation factor (VIF) was computed after each model to check for multicollinearity. All VIFs for the predictors were smaller than 2.00, indicating there was no significant correlation among the independent variables. In other words, there was little overlap of measurements among seven types of learning activity. We opted to report unstandardized coefficients because all the explanatory variables were measured in the same unit (hours). Thus, it was more informative to report the original metrics. Our analysis and visualisations were performed using Stata 13 and Tableau 10.1 respectively.

In line with recent recommendations (Creswell & Plano-Clark, 2011), we employed a mixed-method research design. We applied fixed effect models to a wide range of undergraduate modules at the OU, followed by a triangulated case study approach using six selected first-year modules. Yin (2009) emphasised that a case study investigates a phenomenon in depth and in its natural context. The purpose of a case study is to gain in-depth information about what is happening, why it is happening, and the effects of what is happening.

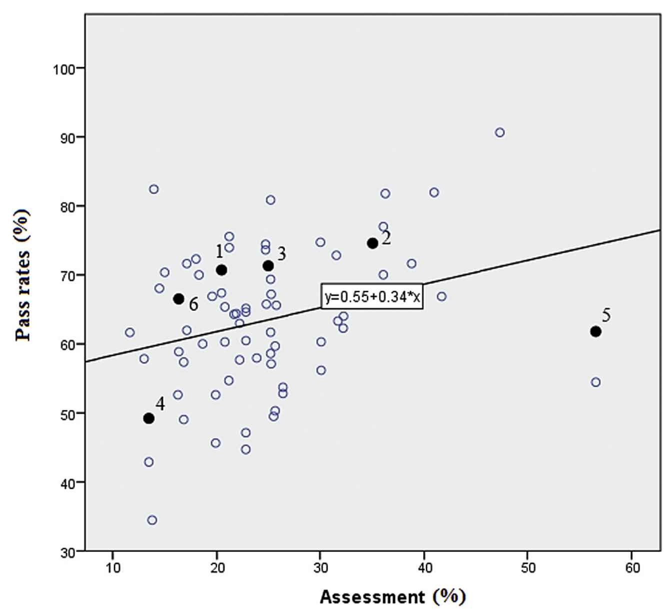

Fig. 1. Scatterplot of the 74 modules showing percentage pass rate of registered students who started the module in relation to the percentage of study time allocated for assessment activities. Note: The six case study modules are marked. R-Squared = 0.097.

Table 2
Descriptive statistics of key variables of 74 modules.

Variable	Obs.	Mean	Std. Dev.	Min	Max
Assimilative (in hours)	74	107.14	60.73	6.00	299.90
Information	74	8.34	12.42	0.00	88.50
Communication	74	8.01	9.65	0.00	43.83
Productive	74	33.47	28.73	0.00	131.60
Experiential	74	8.19	21.78	0.00	164.00
Interactive	74	4.20	8.77	0.00	51.15
Assessment	74	55.48	30.67	17.00	174.25
Total workload	74	224.83	95.43	90.60	515.30
Average time per week (in minutes)	74	90.85	54.58	0.00	304.50
Average time per visit	74	18.33	7.46	0.00	36.93
Registration (no. of students)	74	978.07	824.20	205.00	3707.00
Pass rate (%)	74	63.58	10.22	34.47	90.64
Satisfaction	74	81.22	10.98	38.96	97.34

Note: data aggregated on module level.

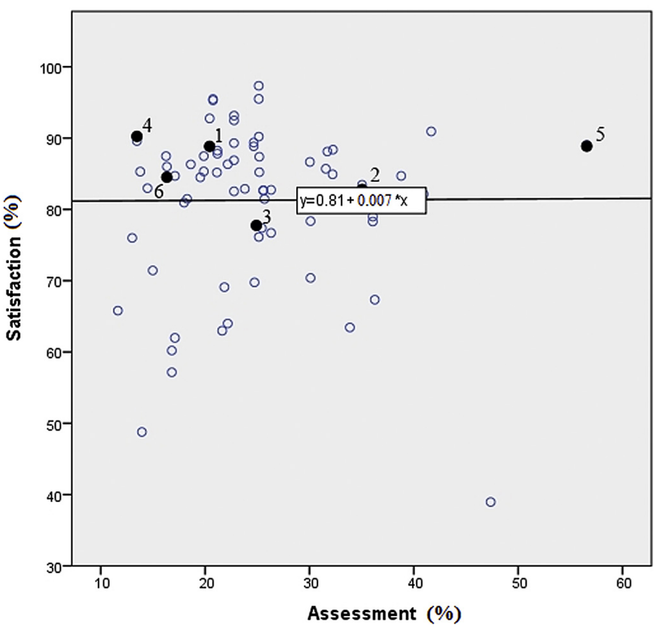

Fig. 2. Scatterplot of the 74 modules showing percentage of learner satisfaction, in relation to the percentage of study time allocated for assessment activities. Note: The six case study modules are marked. R-Squared = 0.000.

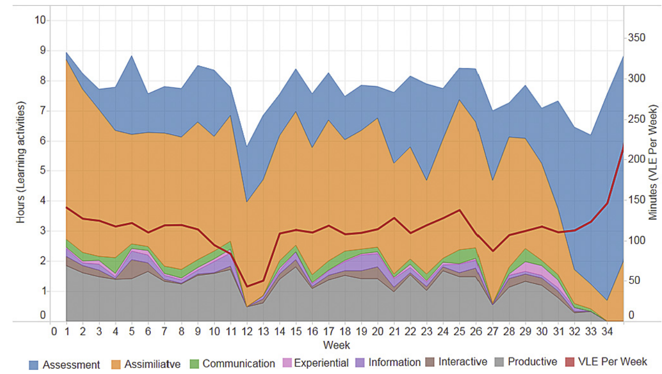

Fig. 3. Longitudinal visualisation of hours per week allocated for different activities in the learning design (coloured blocks) and students' average engagement in minutes per week on the VLE (red line) for 37 modules over 34 weeks. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Using visual inspection of Fig. 1 and Fig. 2, we selected six first-year modules (three Arts/Social Science and three Science/Mathematics) as case studies. Each of the six represented a group of modules within the scatterplots in terms of intensity of assessment and relative performance in terms of pass rates. A detailed review of the relevant learning design data for these modules (balance of activities, assessment types, assessment spread and overall workload) was used to enhance the quantitative analysis of these learning designs.

4. Results

4.1. Assessment design and learning design principles

Research question 1 requires exploration of the extent to which educators used individual CBA designs and how these related to the broader learning design principles of Conole (2012). Descriptive statistics (Table 2) indicated that on average, students were allocated 55.48 h (SD = 30.67) for assessment activities in each module. This accounted for 25.20% of the total workload, with a range of 17.00–174.25 h (Table 2). On the 37 modules for which weekly learning design data were available, students' total workload was on average 7.80 h per week, of which 1.90 h were allocated for assessment, 3.90 h for assimilative activities, 0.20 h for finding information, 0.23 h for communication, 1.30 h for productive activities, 0.08 h for experiential activities, and 0.17 h for interactive activities.

To further unpack the complexity of CBA design on a week-by-week basis, we visualized the seven types of learning activity in 37 modules over 34 weeks (Fig. 3). At a glance, these modules included relatively few assessment activities (blue) in the first three weeks, whereas more assessment activities were assigned at the end of the module, in particular from week 29 onwards where end-of-module assessment was the dominant learning activity. Peaks in assessment can be seen in week 5, in week 10 (before Christmas), in week 23 (before Easter), and at the end of each module.

Bearing in mind how assessment activities had been designed, we went on to examine how CBA designs related to other learning activities. In Table 3, correlation analyses on a weekly level indicate that assessment activities were significantly and negatively correlated with assimilative, communication, information, productive, and experiential activities. To examine the effects of other learning activities on assessment activities, a fixed effect model with assessment as the dependent variable was conducted (Table 4). Assimilative ($B = -0.70$, $SE = 0.02$, $p < 0.01$), information ($B = -0.68$, $SE = 0.09$, $p < 0.01$), communication ($B = -0.73$, $SE = 0.10$, $p < 0.01$), productive ($B = -0.39$, $SE = 0.05$, $p < 0.01$), and experiential ($B = -0.36$, $SE = 0.15$, $p < 0.05$) activities were negatively associated with assessment activities. In other words, using a week-by-week perspective, a rise in the amount of time allocated for assessment activities was significantly linked to a fall in the time allocated for all other types of activity, except for interactive. These negative relationships imply that educators introduced assessment activities while reduced other activities in order to avoid an overwhelming workload. This is an important finding as previous research using only aggregate data showed limited relations between assessment and other learning activities (Rienties & Toetenel, 2016).

4.2. Impact of assessment design on student engagement, student satisfaction and pass rates

4.2.1. Students' engagement in VLE

After exploring the CBA designs, we started to investigate how assessment design are associated with students' activity in the VLE (RQ2). Fixed effect models were conducted with the average time spent on VLE per visit and per week as dependent variables. For each predictor, four models were applied. First, we ran a normal OLS regression model. Second, we used the fixed effect model to control for the unobserved heterogeneity of time. Third, we controlled for the fixed effect between modules. Finally, we controlled for the fixed effects of both time and modules. Since assimilative activities account for most of the workload, they were set as the baseline. Therefore, the following results should be interpreted relative to assimilative activities.

Table 5 shows that assessment activities were positively and significantly related to the average time spent in the VLE per week in all four models. In Models 1 and 2, the effect of assessment activities was almost the same ($B = 4.98$, $SE = 0.57$, $p < 0.01$ and $B = 5.09$, $SE = 0.59$, $p < 0.01$ respectively). The effect of assessment activities became smaller in Model 3 and Model 4 when differences between modules were taken into account. On average, an

Table 3
Correlation matrix of learning design and VLE engagement at weekly level.

Variables	1	2	3	4	5	6	7	8
1. Assessment								
2. Assimilative	−0.46**							
3. Communication	−0.12**	0.17**						
4. Information	−0.12**	0.08**	0.17**					
5. Productive	−0.29**	0.16**	0.13**	0.17**				
6. Experiential	−0.06*	0.02	−0.02	−0.02	0.00			
7. Interactive	0.00	0.02	0.05	0.01	0.01	0.01		
8. VLE per week	0.20**	0.01	0.27**	0.05	0.01	0.01	0.16**	
9. VLE per visit	0.12**	0.10**	0.22**	0.04	0.07*	0.04	0.09**	0.84**

N = 37 modules (1088 data points).
*p < 0.05, **p < 0.01.

Table 4
Fixed effect model of assessment and other learning activities.

DV	(1) Assessment
Assimilative	−0.70** (0.02)
Information	−0.68** (0.09)
Communication	−0.73** (0.10)
Productive	−0.39** (0.05)
Experiential	−0.36* (0.15)
Interactive	0.14 (0.09)
Constant	5.50** (0.12)
Observations	1088
Adjusted R-squared	0.59

Unstandardized coefficients
Standard errors in parentheses.
*p < 0.05, **p < 0.01.

additional hour allocated for assessment activities was associated with 2.47 (SE = 0.47, $p < 0.01$) and 2.80 (SE = 0.47, $p < 0.01$) minutes increase in the average time spent on the VLE per week in Model 3 and Model 4 respectively. Other strong predictors of student engagement in the VLE were communication in all four models and interactive activities in Models 1 and 2. Overall, learning design activities explained up to 58% of the variability in student engagement in the VLE per week when controlling for the heterogeneity within and between modules.

In terms of time spent on the VLE per visit (Table 6), assessment, productive, and experiential activities had strong and positive effects in Models 1 and 2, but became insignificant in Models 3 and 4. Model 2 implied that an additional hour allocated for assessment activities was, on average, associated with a 0.48 min increase in the time spent in the VLE per visit (SE = 0.07, $p < 0.01$). However, the effect became insignificant when controlling for the differences between modules. Additionally, communication activities were positively associated with time on VLE per visit in all models, while productive, experiential, and interactive activities had a significant effect in Models 1 and 2 only. Overall, by taking into account the heterogeneity within and between modules, learning design was able to explain 69% of the variability in time spent on the VLE per visit (Model 4).

4.2.2. *Pass rates and students' satisfaction*

Having identified the effects of CBA designs on VLE engagement, we continued to investigate how assessment activities influence pass rates and students' satisfaction. Scatterplots were created that plotted assessment activities against pass rates (Fig. 1) and against learning satisfaction (Fig. 2). The figures revealed a diversity of approaches to assessment. On average, 25% of students' workload was allocated for assessment activities. However, this average represents a broad range. Some modules, such as Case Study 6 (CS6) and CS4 (see section 4.3 for detailed description of these modules), had a relatively low reliance on assessment activities. Some, like CS5 (which allocated 57% of the total workload for assessment), had a strong focus on assessment activities. Other modules, like CS1 and CS3, were nested around the average of 25%, while CS2 was one standard deviation above the mean. Follow-up linear regressions (Table 7) indicated that CBA assessment activities were positively related to pass rates, while there was no significant relation between assessment activities and satisfaction. Student satisfaction was primarily predicted negatively by student-active learning activities, namely experiential, productive, finding information, and communication. In line with Rienties and Toetenel (2016), student

Table 5
Fixed effect model of VLE engagement per week predicted by learning design activities.

DV = VLE per week	Unstandardized coefficients			
Models	(1)	(2)	(3)	(4)
	OLS	FE_week	FE_module	FE_module_week
Assessment	4.98** (0.57)	5.09** (0.59)	2.47** (0.47)	2.80** (0.47)
Information	2.40 (2.64)	3.23 (2.60)	−0.72 (1.98)	0.15 (1.94)
Communication	26.29** (2.66)	26.29** (2.62)	16.54** (2.16)	17.44** (2.11)
Productive	1.75 (1.14)	1.73 (1.12)	−1.84 (1.04)	−1.83 (1.03)
Experiential	3.57 (3.83)	4.49 (3.78)	−2.07 (2.98)	−0.99 (2.91)
Interactive	11.57** (2.23)	11.25** (2.20)	−0.33 (1.81)	−0.46 (1.78)
Constant	95.66** (2.91)	95.30** (2.85)	110.6** (2.46)	172.1** (10.50)
Observations	1088	1088	1088	1088
Adjusted R-squared	0.15	0.19	0.55	0.58

Standard errors in parentheses.
*p < 0.05, **p < 0.01.
Baseline: assimilative.

Table 6
Fixed effect model of VLE engagement per visit predicted by learning design activities.

DV = VLE per visit	Unstandardized coefficients			
Models	(1)	(2)	(3)	(4)
	OLS	FE_week	FE_module	FE_module_week
Assessment	0.46** (0.07)	0.48** (0.08)	0.02 (0.05)	0.05 (0.05)
Information	0.13 (0.35)	0.20 (0.35)	−0.29 (0.22)	−0.21 (0.21)
Communication	2.76** (0.35)	2.78** (0.35)	0.96** (0.24)	1.06** (0.23)
Productive	0.46** (0.15)	0.46** (0.15)	−0.17 (0.11)	−0.16 (0.11)
Experiential	1.04* (0.51)	1.09* (0.51)	0.52 (0.33)	0.60 (0.32)
Interactive	0.79** (0.30)	0.72* (0.30)	−0.34 (0.20)	−0.39 (0.20)
Constant	20.56** (0.39)	24.15** (1.37)	21.00** (0.91)	24.66** (1.18)
Observations	1088	1088	1088	1088
Adjusted R-squared	0.08	0.10	0.67	0.69

Standard errors in parentheses.
*p < 0.05, **p < 0.01Baseline: assimilative.

Table 7
Linear regressions of student satisfaction and pass-rates by learning design activities.

DV	Unstandardized coefficients	
	(1)	(2)
	Satisfaction	Pass rates
Assessment	−0.09 (0.13)	0.30* (0.12)
Information	−0.52* (0.25)	0.32 (0.26)
Communication	−0.66* (0.33)	0.38 (0.35)
Productive	−0.32** (0.10)	0.04 (0.10)
Experiential	−0.59** (0.20)	−0.23 (0.20)
Interactive	0.37 (0.31)	0.23 (0.32)
Constant	0.94** (0.04)	0.54** (0.05)
Observations	74	74
Adjusted R-squared	0.29	0.11

Standard errors in parentheses.
*p < 0.05, **p < 0.01.
Baseline: assimilative.

satisfaction was positively related to assimilative learning activities (i.e., the benchmark variable).

4.3. *Assessment strategies in relation with student engagement over time*

After carrying out the quantitative analyses, we selected six case study modules based on their positions in Figs. 1 and 2. Additional fine-grained qualitative (e.g., assessment brief for the assessment task) learning design data were available for these modules. For each of these case study modules, the balance of activities was considered to see how the time spent on assessment compared to the balance of other activities. Findings are summarised in the assessment strategy column in Table 8.

As Table 8 outlines, there were substantial variations in the main assessment methods across the six case studies (i.e., assignment, face-to-face exam, final assignment and interactive computer-marked assessment). However, assignments were the main form of assessment in each of the six case studies. In most cases, students were asked to answer a number of questions and then write a short essay. Two of the six case studies (CS4, CS5) included interactive CBAs, which tested students' factual knowledge. These modules included three to seven assignments.

Overall, the assessment approach at the OU is to minimise the use of exams at this level (the equivalent of first-year undergraduate), where possible, as institutional data suggest that the use of final assignments is related to both satisfaction and pass rates (Li et al., 2016). It is therefore interesting that CS3 (Fig. 5) and CS4 (Fig. 6) both used an exam but their respective pass and satisfaction varied widely. Furthermore, CS4 included practice quizzes to help students prepare for their interactive CBAs, whilst CS3 included assessment preparation tasks for each of the assignments. The only module in which assessment took place almost continuously was CS4, which performed well in terms of satisfaction but not in terms of student pass rates. However, when looking at the learning design data in more detail, each case study module included 6–28 assessment activities. This is an interesting finding, as many educators would expect to see at least one assessment activity per week in courses of this length.

Figs. 4–6 illustrate the learning design of the three case studies over 34 weeks. At a glance, assessment and assimilative activities were the dominant designs. As was evident in Fig. 3, there were gaps in weeks 12–13 and week 24 due to Christmas and Easter breaks, as well as breaks in the schedule to prepare for the next learning activities. Assessments were typically conducted every three weeks throughout the course, except in the case of CS4 (Fig. 6). For CS4 a continuous line of assessment can be seen (the blue block) because students were working towards four assessments. In line with our fixed effect analyses, teachers seemed to include assessment at the expense of other learning activities.

A positive relation between assessment and time spent on VLE per week is reflected in Figs. 4–6, as VLE engagement usually went up in weeks with assessment activities. The workload fluctuated considerably between assessment weeks and non-assessment weeks. For example, in CS3 (Fig. 5) the workload in the assessment weeks was similar to that in non-assessment weeks. However, in CS2 (Fig. 4) students were expected to study for more time during assessment weeks than in non-assessment weeks. In other words, both the qualitative narrative, the time dedicated to assessment activities, and actual student behaviour highlight substantial differences in the way these six modules supported learners and provided assessment for and of learning.

5. Discussion

Assessment and feedback are key drivers for learning (Conole, 2012; Scherer et al., 2014; Segers et al., 2003). With the omnipresence of CBA in blended, distance learning and MOOCs, one would expect that some consistent patterns of design of assessment across modules on a macro level could be identified. Pedagogy and learning design are of key importance in online learning (Conole, 2012; Kirschner et al., 2004) but limited research has made explicit links between learning design, learning behaviour, and

Table 8
Overview of assessment strategies of case study modules.

	Context	Assessment strategy	CBA example
CS1	This introductory Arts module focuses on the past and present concerns of Arts disciplines.	Assessment is estimated to take 20% of the directed study time allocated to this module and takes place every 3–4 weeks. There are seven assignments and one final assignment at the end of the module.	All assignment briefs are provided online and assignments are submitted online. Assignments are hand marked by tutors, but the personalised feedback is provided online when the assignment is returned.
CS2	This introductory practical module focuses on personal finance.	More than a third of directed study time is allocated for assessment, which takes place every four weeks. An online tutorial is provided every two weeks.	Assignments are the main form of assessment in this module. Learners are also provided with an online checklist to ensure they have included all detail required.
CS3	This interdisciplinary module focuses on the broad concepts of voices, texts and material culture.	A quarter of the time in this module is allocated for assessment. There are six assignments in addition to a face-to-face exam. For each of the assignments, a preparation task is provided.	Assignments are the main form of assessment in this module. Learners are provided with an assignment preparation task for each assignment.
CS4	This module provides a broad foundation for university-level Mathematics as well as preparation for subjects such as Physics, Engineering and Economics.	Assessment is continuous in the first 20 weeks of the module, which then focuses on preparation for the face-to-face exam. There are four assignments and four interactive computer-marked assessments in this module.	Assignments are the main form of assessment in this module. Four interactive computer-marked assessments are used to assess factual knowledge.
CS5	This introductory module develops knowledge and skills required for Engineering and prepares students for further study in Engineering or a related subject	Assessment takes place every 3–4 weeks and includes four assignments, interactive computer-marked assessments and an end-of-module assignment.	Assignments are the main form of assessment in this module.
CS6	This introductory module develops an understanding of design. Students acquire new design skills and build a portfolio of design projects.	Assessment takes up much of the time allocated to directed study for this module and is spread throughout the module. The time allocated for assessment is high because students are encouraged to work on their assignment for several weeks prior to submission.	Assignments are the main form of assessment in this module. This module also uses peer feedback methods. Students share their work and comment on this online. Participation in this process is a small part of the assessment mix for the module.

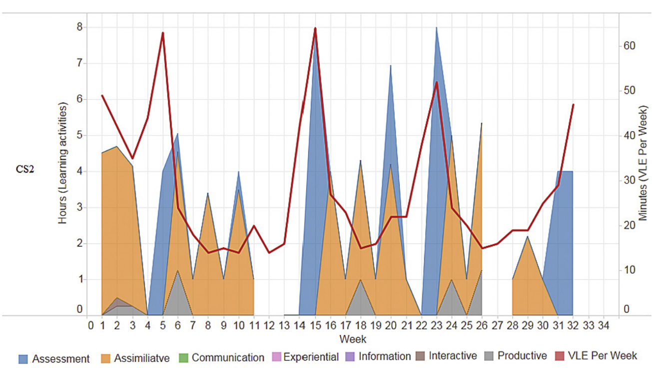

Fig. 4. Longitudinal visualisation of learning design (coloured blocks) and average students' engagement (red line) in the VLE each week for CS2. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

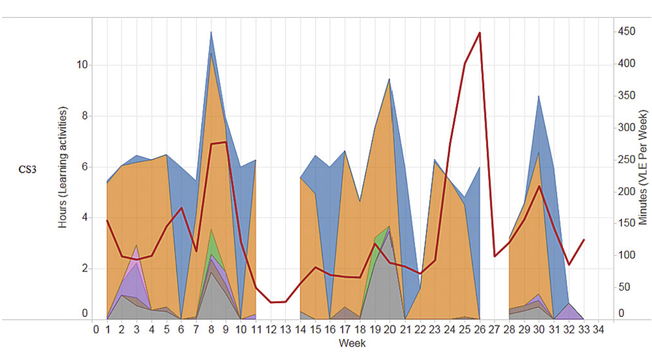

Fig. 5. Longitudinal visualisation of learning design (coloured blocks) and average students' engagement (red line) in the VLE each week for CS3. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

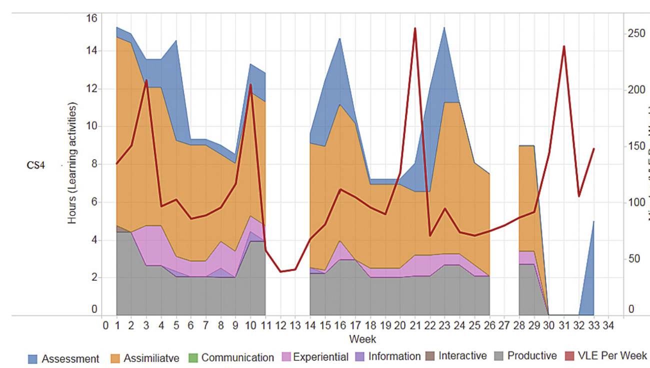

Fig. 6. Longitudinal visualisation of learning design (coloured blocks) and average students' engagement (red line) in the VLE each week for CS4.

learner performance due to a lack of integrated, linked empirical data (Kirschner, Sweller, & Clark, 2006; Rienties & Toetenel, 2016; Toetenel & Rienties, 2016). In the UK, the OU has been leading developments in data gathering through a comprehensive learning design system, which is used by specialists in order to code learning activities. This means that time spent on a range of learning activities can be quantified and analysed in an empirical manner.

Our first, and perhaps most interesting, finding is that although time allocated for assessment varied between 13% and 57% of overall study time, CBA seemed to be balanced with other learning design activities (research question 1). As was previously found on an aggregate level (Rienties & Toetenel, 2016), teachers tended to balance assessment activities with other learning design activities, as our fixed effect model (Table 4) indicated negative betas for all six learning design activities. In other words, in line with good assessment practice (Boud & Falchikov, 2006; Koedinger et al., 2013; Segers et al., 2003) on a weekly level educators did seem to balance assessment with the other six learning design activities, rather than using it to increase student workload in those weeks.

Our second important finding was that the design of weekly learning design activities significantly predicted VLE engagement by students, explaining up to 69% of variance (research question 2). While assessment did significantly predict students' online behaviour over time, communication and interactive activities also helped to explain variations in VLE engagement within and across modules. Furthermore, the relative frequency of time allocated for assessment activities was significantly related to student pass rates, but no clear relation with satisfaction was found, in line with previous research (Rienties & Toetenel, 2016). Recent research has highlighted a need for learning analytics researchers to take time into consideration (Greiff et al., 2016; Kovanovic et al., 2015; Tempelaar et al., 2015).

Our findings indicate that VLE engagement is not only substantially influenced by learning design, but is particularly influenced by how educators within and across modules balance their learning design activities on a week-by-week basis. Building on conceptual work related to learning design (Armellini & Aiyegbayo, 2010; Conole, 2012; Lockyer et al., 2013; Toetenel & Rienties, 2016), this study provides further empirical evidence that, in order to understand how students are learning online and how to analyse fluctuations in their online behaviour over time, learning design should play a central role in LA modelling in general, and CBA modelling in particular.

Finally, our six case studies highlight the complexities that educators face when developing appropriate learning designs (research question 3). Very different sets of learning activities were built into the designs of these six modules, and these designs significantly influenced how students engaged with the VLE (see Figs. 4–6). Two case studies (CS4 and CS5) incorporated assessment throughout their module, in line with the learning outcomes. This is considered to be good practice: teaching and assessment should be aligned. However, these modules did not perform better than the modules in which assessment activities were related to more defined assessment points in the module.

Our fine-grained analyses of these six case studies showed that the modules had very different assessment strategies. This was a surprise, given that extensive production and validation processes are in place at the OU. When students are confronted with different assessment strategies with each new module they take, this requires them to adapt their learning strategies (Bearman et al., 2016; Boud & Falchikov, 2006). While some variation in learning design and particularly in assessment may encourage interest and motivation for some students, others with less flexible learning strategies may struggle to cope if a previously successful learning strategy is not effective for the next assessment (Tempelaar et al., 2017 ; Terzis & Economides, 2011; Wolff, Zdrahal, Nikolov, & Pantucek, 2013).

6. Limitations and future work

One obvious limitation of our research is that we did not include more micro level fine-grained analyses of actual student behaviour on CBA tasks, such as those done by Greiff et al. (2016). Given the size of its student population, the OU is currently only able to keep fine-grained data for the modules on which predictive analytics engines are being piloted instead of individual student data (Wolff et al., 2013). In the near future, we aim to analyse the types of strategy that individual students employ when dealing with different CBA designs. This will help to link our macro-understanding of learning design and assessment with fine-grained individual user behaviour.

A second limitation is that we used rather crude measurements of course success (pass rates and satisfaction). Whether students were actually learning and to what extent good pass rates and high satisfaction scores are actually an accurate reflection of appropriate learning designs and CBA can be debated. The diversity of disciplines and students on Level 1 modules (Rienties & Toetenel, 2016) can influence how satisfaction is interpreted and measured, which undermines the comparability of course success across modules. As a consequence of lacking individual student data, we were not able to perform a multiple-group confirmatory factor analysis across modules to address the issue of measurement invariance of the latent construct, satisfaction. Multi-level models taking into consideration both learner characteristics (Tempelaar et al., 2015; Terzis & Economides, 2011) and learning designs (Hernández-Leo et al., 2014; Lockyer et al., 2013; Toetenel & Rienties, 2016) will need to be explored in the future in order to determine which CBA designs are appropriate for which type of learner and discipline.

Further longitudinal research following students over different modules across a qualification is needed in order to understand how flexible and effective learners are in adjusting their learning strategies to the assessment strategy of a particular module, and whether there is a need to harmonise assessment designs across modules. It is clear from the fine-grained learning design analysis undertaken that further research into the use of formative assessment and ways of coding this information would be helpful in order to understand the role of assessment spread, load and forms of CBA and their impact on student outcomes and satisfaction.

7. Conclusion

This study examined the design of CBA and its effects on student engagement in the VLE, satisfaction, and pass rates using a combination of visualisations, and mixed methods on 74 first-level undergraduate modules at the OU. The first finding indicated that, on average, assessment activities accounted for 25% of the total workload, with great variability across modules. Moreover, the workload on other activities decreased when assessment activities were introduced. This implied that educators aimed to balance the total workload when designing CBA.

Secondly, assessment activities were associated with more time spent in the VLE, since the majority of assessment activities required computer use. Learning design in general could explain up to 69% of the variance in VLE engagement. Modules with higher relative frequency of assessment activities are associated with higher pass rates, while no clear relation was found with regard to satisfaction.

Finally, the six case study modules illustrated the diversity of assessment strategies across disciplines in terms of assessment types and approaches. Further visualisations support our findings above as educators reduced the workload on other activities when introducing assessment, and the time spent in VLE increased considerably in assessment weeks.

Our study contributes to CBA innovation by approaching CBA from a learning design perspective using learning analytics. While CBA has been studied extensively in terms of micro-analysis on individual assessment items (Greiff et al., 2015, 2016) and assessment strategies within a module (Brito & de Sá-Soares, 2014; Tempelaar et al., 2015; Terzis & Economides, 2011), the connection between CBAs and learning designs across a large number of distance learning modules has not been explicitly examined in the past. Therefore, by investigating CBA and other learning activities in tandem, we provided a broader picture of how teachers design assessment activities within their modules and across time.

In line with recent reviews of learning analytics (Ferguson et al., 2016 ; Papamitsiou & Economides, 2014, 2016), we encourage researchers to look beyond “cold” learning analytics data (such as weekly learning design activities, CBAs and student engagement), as a rich diversity of practice seemed to be present when contrasting six case studies. While our findings highlight a strong relation between weekly learning design activities and student engagement, it is important to bring teachers and educational researchers on board to unpack the complexities of learning.

In terms of practical implications, assessment and feedback are high on the priority list for students and educators, as these link directly to student success and to the success of a course, programme, faculty and university. Some policy makers have already made moves intended to improve the effectiveness of teaching (Ferguson et al., 2016). For example, a Teaching Excellence Framework has been introduced in the UK, and it is likely that measures related to assessment will be used as key indicators. In order to explain how satisfaction and assessment activities are linked and which elements of assessment (balance of activities, spread through module material or assessment methods) have a significant impact on student outcomes, we need to combine research data and institutional data and work together in order to solve this complex puzzle.

Acknowledgements

We would like to thank the reviewers for their excellent and detailed feedback. In particular, we are grateful for the continuous encouragements from the special editor of this special issue on Innovations in CBA, Prof Samuel Greiff Quan Nguyen is supported and funded by the Leverhulme Trust Research Project Grant "Open World Learning".

References

Agudo-Peregrina, Á. F., Iglesias-Pradas, S., Conde-González, M.Á., & Hernández-García, Á. (2014). Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning. *Computers in Human Behavior, 31*(February), 542–550. http://dx.doi.org/10.1016/j.chb.2013.05.031.

Arbaugh, J. B. (2014). System, scholar, or students? Which most influences online MBA course effectiveness? *Journal of Computer Assisted Learning, 30*(4), 349–362. http://dx.doi.org/10.1111/jcal.12048.

Armellini, A., & Aiyegbayo, O. (2010). Learning design and assessment with e-tivities. *British Journal of Educational Technology, 41*(6), 922–935. http://dx.doi.org/10.1111/j.1467-8535.2009.01013.x.

Bearman, M., Dawson, P., Boud, D., Bennett, S., Hall, M., & Molloy, E. (2016). Support for assessment practice: Developing the assessment design decisions framework. *Teaching in Higher Education, 21*(5), 545–556. http://dx.doi.org/10.1080/13562517.2016.1160217.

Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education, 31*(4), 399–413. http://dx.doi.org/10.1080/02602930600679050.

Brito, M. A., & de Sá-Soares, F. (2014). Assessment frequency in introductory computer programming disciplines. *Computers in Human Behavior, 30*, 623–628. http://dx.doi.org/10.1016/j.chb.2013.07.044.

Conole, G. (2012). *Designing for learning in an open world*. Dordrecht: Springer.

Creswell, J. W., & Plano-Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). London: Sage Publications.

Cross, S., Galley, R., Brasher, A., & Weller, M. (2012). *Final project report of the OULDI-JISC project: Challenge and change in curriculum design process, communities, visualisation and practice*. York: JISC.

Ferguson, R., Brasher, A., Cooper, A., Hillaire, G., Mittelmeier, J., Rienties, B., … Vuorikari, R. (2016). Research evidence of the use of learing analytics; implications for education policy. In R. Vuorikari, & J. Castano-Munoz (Eds.), *A European Framework for Action on Learning Analytics*. Luxembourg: Joint Research Centre Science for Policy Report.

Ferguson, R., & Buckingham Shum, S. (2012). Social learning analytics: five approaches. In *Paper presented at the 2nd international conference on learning analytics and knowledge, Vancouver, British Columbia, Canada*.

Greiff, S., Niepel, C., Scherer, R., & Martin, R. (2016). Understanding students' performance in a computer-based assessment of complex problem solving: An analysis of behavioral data from computer-generated log files. *Computers in Human Behavior, 61*, 36–46. http://dx.doi.org/10.1016/j.chb.2016.02.095.

Greiff, S., Wüstenberg, S., & Avvisati, F. (2015). Computer-generated log-file analyses as a window into students' minds? A showcase study based on the PISA 2012 assessment of problem solving. *Computers & Education, 91*, 92–105. http://dx.doi.org/10.1016/j.compedu.2015.10.018.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica, 46*(6), 1251–1271. http://dx.doi.org/10.2307/1913827.

Hernández-Leo, D., Moreno, P., Chacón, J., & Blat, J. (2014). LdShake support for team-based learning design. *Computers in Human Behavior, 37*, 402–412. http://dx.doi.org/10.1016/j.chb.2012.05.029.

Kapur, M. (2008). Productive failure. *Cognition and Instruction, 26*(3), 379–424. http://dx.doi.org/10.1080/07370000802212669.

Kirschner, P. A., Strijbos, J.-W., Kreijns, K., & Beers, P. (2004). Designing electronic collaborative learning environments. *Educational Technology Research and Development, 52*(3), 47–66.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86. http://dx.doi.org/10.1207/s15326985ep4102_1.

Koedinger, K., Booth, J. L., & Klahr, D. (2013). Instructional complexity and the science to constrain it. *Science, 342*(6161), 935–937. http://dx.doi.org/10.1126/science.1238056.

Kovanovic, V., Gasevic, D., Dawson, S., Joksimovic, S., Baker, R. S., & Hatala, M. (2015). Penetrating the black box of time-on-task estimation. In *Paper presented at the 5th learning analytics knowledge conference, Poughkeepsie*.

Li, N., Marsh, V., Rienties, B., & Whitelock, D. (2016). Online learning experiences of new versus continuing learners: A large scale replication study. *Assessment & Evaluation in Higher Education*. http://dx.doi.org/10.1080/02602938.2016.1176989.

Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing pedagogical action: Aligning learning analytics with learning design. *American Behavioral Scientist, 57*(10), 1439–1459. http://dx.doi.org/10.1177/0002764213479367.

MacLean, P., & Scott, B. (2011). Competencies for learning design: A review of the literature and a proposed framework. *British Journal of Educational Technology, 42*(4), 557–572. http://dx.doi.org/10.1111/j.1467-8535.2010.01090.x.

Moreno-Ger, P., Burgos, D., Martínez-Ortiz, I., Sierra, J. L., & Fernández-Manjón, B. (2008). Educational game design for online education. *Computers in Human Behavior, 24*(6), 2530–2540. http://dx.doi.org/10.1016/j.chb.2008.03.012.

Onwuegbuzie, A. J., Witcher, A. E., Collins, K. M. T., Filer, J. D., Wiedmaier, C. D., & Moore, C. W. (2007). Students' perceptions of characteristics of effective college teachers: A validity study of a teaching evaluation form using a mixed-methods analysis. *American Educational Research Journal, 44*(1), 113–160. http://dx.doi.org/10.3102/0002831206298169.

Open University UK. (2016). *Assessment handbook (A. P. Office, Trans.)*. Milton Keynes: Open University UK.

Papamitsiou, Z., & Economides, A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Educational Technology & Society, 17*(4), 49–64.

Papamitsiou, Z., & Economides, A. (2016). Learning analytics for smart learning environments: A meta-analysis of empirical research results from 2009 to 2015. In J. M. Spector, B. B. Lockee, & D. M. Childress (Eds.), *Learning, design, and Technology: An international compendium of theory, research, practice, and policy* (pp. 1–23). Cham: Springer International Publishing.

Ras, E., Whitelock, D., & Kalz, M. (2015). The promise and potential of e-assessment for learning. In P. Reimann, S. Bull, M. Kickmeier-Rust, R. Vatrapu, & B. Wasson (Eds.), *Measuring and visualizing learning in the information-rich classroom* (pp. 21–40). New York: Routledge.

Rienties, B., & Toetenel, L. (2016). The impact of learning design on student behaviour, satisfaction and performance: a cross-institutional comparison across 151 modules. *Computers in Human Behavior, 60*, 333–341. http://dx.doi.org/10.1016/j.chb.2016.02.074.

Rogaten, J., Rienties, B., & Whitelock, D. (2016). Assessing learning gains. In *Paper presented at the technology enhanced assessment conference 2016, Tallinn, Estonia*. http://oro.open.ac.uk/47884/.

Scherer, R., Meßinger-Koppelt, J., & Tiemann, R. (2014). Developing a computer-based assessment of complex problem solving in Chemistry. *International Journal of STEM Education, 1*(1), 1.

Segers, M., Dochy, F., & Cascallar, E. (2003). *Optimising new modes of assessment: In search of qualities and standards*. Dordrecht: Kluwer Academic Publishers.

Tempelaar, D. T., Rienties, B., & Giesbers, B. (2015). In search for the most informative data for feedback generation: Learning analytics in a data-rich context. *Computers in Human Behavior, 47*, 157–167. http://dx.doi.org/10.1016/j.chb.2014.05.038.

Tempelaar, D. T., Rienties, B., & Nguyen, Q. (2017). Towards actionable learning analytics using dispositions. *IEEE Transactions on Learning Technologies, PP*(99). http://dx.doi.org/10.1109/TLT.2017.2662679.

Terzis, V., & Economides, A. A. (2011). Computer based assessment: Gender differences in perceptions and acceptance. *Computers in Human Behavior, 27*(6), 2108–2122. http://dx.doi.org/10.1016/j.chb.2011.06.005.

Toetenel, L., & Rienties, B. (2016). Analysing 157 Learning Designs using Learning Analytic approaches as a means to evaluate the impact of pedagogical decision-making. *British Journal of Educational Technology, 47*(5), 981–992. http://dx.doi.org/10.1111/bjet.12423.

Vandewaetere, M., Desmet, P., & Clarebout, G. (2011). The contribution of learner characteristics in the development of computer-based adaptive learning environments. *Computers in Human Behavior, 27*(1), 118–130. http://dx.doi.org/10.1016/j.chb.2010.07.038.

Wolff, A., Zdrahal, Z., Nikolov, A., & Pantucek, M. (2013). Improving retention: Predicting at-risk students by analysing clicking behaviour in a virtual learning environment. In *Paper presented at the proceedings of the third international conference on learning analytics and knowledge, indianapolis*.

Yin, R. K. (2009). *Case study research: Design and methods* (5 ed.). Thousand Oaks:

Sage.
Zerihun, Z., Beishuizen, J., & Os, W. (2012). Student learning experience as indicator of teaching quality. *Educational Assessment, Evaluation and Accountability, 24*(2), 99–111. http://dx.doi.org/10.1007/s11092-011-9140-4.

Quan Nguyen is a Ph.D. candidate in Learning Analytics, Institute of Educational Technology, The Open University, UK. He graduated in 2016 as a Master (cum laude) of Economics and Information Management at Maastricht University. His research focusses on the use of learning analytics to understand how teachers design their course and how learning design could affect other dimensions of the learning progress such as engagement, satisfaction, or performance

Dr Bart Rienties is Reader in Learning Analytics at The Open University, UK. His research interests include learning analytics, collaborative learning environments and the role of social interaction in learning.

Lisette Toetenel is a Senior Learning & Teaching Development Manager at The Open University, UK. Her research interests range from learning design and learning analytics to social media in education.

Dr Rebecca Ferguson is a senior lecturer in the Institute of Educational Technology (IET) at The Open University in the UK and a senior fellow of the Higher Education Academy. Her primary research interests are educational futures and how people learn together online.

Professor Denise Whitelock's main research interests are concerned with automatic feedback systems that promote assessment for learning. These are generating new thinking about modelling automatic socio-emotive feedback, which combined with automated constructive cognitive comments, can enhance both student learning and motivation.